View metadata, citation and similar papers at core.ac.uk
brought to you by CORE

Differences between English and Japanese Speakers' Use of Metadiscourse to Identify Discourse Units in English Research Article Abstracts

Akiko Hagiwara,[1] Kaoru Kobayashi,[2] Mitsuki Fukuike[3]

Introduction

The research article abstract, or summary, is written discourse that belongs to professional academic communities, and its primary function is to give a compact summary of the article. In the field of life sciences, there are two main types: structured abstracts and unstructured abstracts (American Medical Association, 2016). Structured abstracts, used in many clinical health journals such as *The New England Journal of Medicine* and *The Lancet*, are those with labeled sections: 'Objectives,' 'Methods,' 'Results,' and 'Conclusion.' Other journals have unstructured abstracts consisting of a single paragraph with similar pieces of information. This type is used in major journals that focus on fundamental research in biology, e.g., *Cell*, *Nature*, *Genetics*. Compared to structured abstracts, reading unstructured abstracts is linguistically more demanding because of the lack of headings. Native speakers of English (NS) may be able to follow the logical development without much difficulty, but for second language readers (NNS), recognizing logical structure may not be easy because the information units in the text are not clearly marked, especially when their content knowledge is limited. In order to successfully understand an unstructured abstract, it is important for an L2 reader to develop a strategy to identify each unit of information.

Ample research on unstructured abstracts can be found in genre analysis and across disciplines such as second language acquisition studies and in the field of teaching English to speakers of other languages (TESOL). Many studies use the terms 'moves' and 'steps' as units of analysis (Bhatia, 1993; Samraj, 2004; Cross and Oppenheim, 2006; Stoller and Robinson, 2013). A move consists of one or more sentences about a particular topic and is a component of a particular section of a genre, such as the introduction section of a research article. Connor and Mauranen (1999) define a move as a functional unit that contains one or more propositions, has a clear rhetorical purpose, and can be further divided into 'steps.' In general, the abstract of a research article consists of four moves (Bhatia 1993): introducing purpose, describing methodology, summarizing results, and presenting conclusions. Thus, the structure of research article abstracts somewhat resembles the structure of research articles. This suggests that, for junior researchers whose native language is not English, being able to read the abstract may help them to fully understand the entire research article.

[1] 生命科学部言語科学研究室
[2] 生命科学部非常勤講師
[3] 生命科学部言語科学研究室

One strategy is to learn about the general rhetorical organization of unstructured abstracts, but it is not an easy task, because research journals have style preferences. Research shows that the rhetorical structure of research article abstracts varies slightly from discipline to discipline. Based on their study of 12 scientific abstracts in the field of protozoology, Cross and Oppenheim (2006) claim there are five moves in research article abstracts: relation to other research, purpose, methodology, summarizing the results, and discussing the research. Stoller and Robinson (2013) explain that chemistry journal abstracts often start with a statement of what was done, followed by identification of methods, and a report of principal findings. These findings suggest that each discipline or even each journal within a discipline, in science may have a specific style. Biology journals including *Nature*, *Cell*, and *Science* have specific guidelines for constructing the abstract/summary section with explanations on what information is to be included and in what order. For example, *Nature*, gives the author specific guidelines to start with 'a basic introduction to the field' followed by 'more detailed background,' a statement of 'the general problem,' the introduction of the main result followed by an explanation, an explanation on how to put the results in a 'general context,' and finally, a broader perspective on how the results can be utilized (Nature, 2016). Sometimes a journal may even give the author specific advice on how many sentences they should write on each of the above topics. Thus, learning the style of each journal may not be easy for NNS researchers.

A better strategy for NNS researchers may be to learn different types of linguistic or paralinguistic devices that signal the boundaries of moves and steps. In written discourse, linguistic devices such as conjunctions are often used to mark the boundaries of moves. Connor and Mauranen (1999) claim that discourse markers such as *consequently, however, firstly, to sum up* are among the devices used to identify moves. Hyland and Tse (2004) call them interactive resources of metadiscourse that "help to guide the reader through the text" (p. 169) and further categorize them into five groups (Table 1). Among these, transition markers, such as *however*, *the result is*, and frame markers, such as *to summarize*, *in conclusion*, are considered to be signaling the shift of moves and steps.

Table 1

Interactive Metadiscourse

Category	*Functions*	*Examples*
Transitions	express semantic relation b/w main clauses	in addition/but/thus/and
Frame m.	refer to discourse sequences, or text stages	finally/to conclude/my purpose is to
Endophoric m.	refer to information in other parts of the text	noted above/see Fig/in section 2
Evidentials	refer to source of info. from other texts	according to X/(Y,1990)/Z states
Code glosses	help readers grasp functions of ideational material	namely/e.g./such as/in other words

Note: Adapted from Hyland and Tse (2004, Table 1, rows 1-7, p. 169)

The uses of metadiscourse vary among disciplines and depend on the language (Dahl, 2004; Peacock, 2010; Carrio-Pastor, 2013; Cao and Hu, 2014). Peacock's (2010) study on the

use of linking adverbials shows disciplinary differences in the use of these metadiscourse phrases. 320 research articles, 160 from four disciplines of science (chemistry, computer science, materials science, and neuroscience) and 160 from four non-science disciplines (economics, language and linguistics, management, and psychology) were compared. Four types of adverbials were analyzed: contrast/concession (e.g., *however* and *instead)*, result/inference (e.g., *thus* and *therefore*), addition (e.g., *also* and *as well*), and apposition (e.g., *for example* and *in other words*). A significantly higher occurrence of linking adverbials was present in science articles than in non-science articles. Among the science disciplines, the occurrence rate was lower in chemistry and materials science than in computer science and neuroscience. In chemistry and materials science, Peacock presumes the rhetorical structure is basically fixed and that there is no need to explicitly show the connection between the information units.

Cao and Hu (2014) also showed cross-disciplinary variations, using Hyland and Tse's (2004) five types of interactive metadiscourse (Table 1). Their corpus consisted of 120 quantitative and qualitative research articles, 40 each from education, psychology and applied linguistics disciplines. Cross-disciplinary differences were observed in four categories of interactive metadiscourse but not in frame markers. Applied linguistics articles employed more transitional markers and endophoric markers than psychology articles and more evidential markers than both psychology and education articles.

Dahl (2004) studied the number of occurrences of metadiscourse phrases in a corpus of 180 research articles, 60 written in each of English, French and Norwegian. Of the 60 research articles in each language, 20 each were from linguistics, economics and medicine. The types of metadiscourse studied were locational metatext that points to the text itself or a particular section of it, and rhetorical metatext which is similar to Hyland and Tse's category of frame markers (Table 1). The results showed that the amount of metadiscourse used in the articles written in French was much smaller than that in either English or Norwegian. Among the three disciplines, the metatext phrases in medicine are similar in three languages, but in economics and linguistics they were different. The author argues that in medicine, there is an established rhetorical structure of the genre across languages, but for economics and linguistics, the difference is due to the tradition among the Anglo-Saxons and the Scandinavians where the writers have a responsibility for guiding the readers through the logical development of the text. As for economics and linguistics, the figures for economics in both English and Norwegian were higher than those in linguistics. The results suggest that the use of metadiscourse is dependent both on the discipline and the language.

Carrio-Pastor (2013) studied the use of sentence connectors (i.e., transition markers and frame markers) in two corpora: 20 academic papers in engineering written in English by native speakers of English (NS) and 20 written in English by native speakers of Spanish

(NNS). The results showed that NNS tend to use more connectors than NS. In the same number of running words in each corpus, the total numbers of connectors found in the corpora were 869 (NS) and 1,138 (NNS). The most frequently used categories of listing (e.g., *finally, furthermore,* and *in addition*) and contrast (e.g., *however, although,* and *on the other hand*) were used similarly by both groups, but they rarely used connectors of colloquial style such as *in brief, in short, all in all, in other words, eventually* and *in sum*. While NNS used more summative connectors (e.g., *to sum up*, *to conclude*, and *in summary*), NS used more transitional (e.g., *meanwhile, in the meantime, originally,* etc.), appositional (e.g., *for example, namely, that is*, etc.), resultive (e.g., *consequently, therefore, as a result*, etc.) and inferential connectors (e.g., *therefore, in that case, otherwise*, etc.). The author explains the similarities may come from the tendency of NNS to mimic the way NS write and attributes the differences to the fact that Spanish writers expect the readers to be more responsible for understanding the text than English writers do.

The above studies show that variations in metadiscourse resources are attributed to the specific genre, the writer's linguistic background, and the academic discipline and we can conclude that one strategy for NNS to understand research abstracts is to familiarize themselves with metadiscourse phrases in a given discipline. The objective of the present study is to identify the metadiscourse phrases that NNS researchers should learn in the discipline of life sciences. For this objective, the focus was placed on transitions and frame markers (Table 1) and the following research questions were addressed.

1. What metadiscourse markers are used in abstracts written by NS and NNS?
2. Which transition markers and frame markers are used in abstracts?
3. What are the differences in the use of these metadiscourse markers between abstracts written by NS and those written by NNS?

Corpora

We collected two corpora: a NS corpus with 2000 abstracts from various fields of life sciences presumably written by scientists in English speaking countries; a NNS corpus with 2000 abstracts in the same fields mainly written by Japanese scientists in Japan. We categorized them based on the authors' first and last names and their affiliations.

Results

This section reports on the quantitative analyses obtained in our study.

After collecting the abstracts, frequency word lists were obtained from both NS and NNS corpora with 390,165 running words and 375,113 running words respectively. Table 2 shows the types and tokens of each corpus. Then, we ran an *N*-gram analysis in order to obtain clusters, which yielded a list of two- to five-word sequences that appeared in each corpus.

Table 2

Description of NNS and NS Corpora

Corpus	NS	NNS
Types	11,148	11,083
Tokens	390,165	375,113

After generating two lists of clusters, we matched them with the list of metadiscourse markers classified by Hyland (2005). The obtained results are summarized in Table 3. The results show various metadiscourse markers are used in abstracts, and between the two categories of metadiscourse, interactive metadiscourse markers were used more frequently than interactional markers, and among the interactive markers, transition markers were predominantly used.

Table 3

Types and Tokens of Metadiscourse Markers Identified in NNS and NS Corpora

	Types		Tokens	
Metadiscourse Markers	NS	NNS	NS	NNS
Interactional metadiscourse	*175*	*171*	*18,244*	*16,785*
Attitude Markers	24	24	776	748
Boosters	28	24	952	912
Engagement Markers	61	58	9,143	7,875
Hedges	56	58	4,074	4,260
Self Mention	6	7	3,299	2,990
Interactive metadiscourse	*87*	*88*	*23,940*	*22,550*
Code Glosses	10	11	3,955	3,200
Endophoric Markers	1	1	76	57
Frame Markers	31	31	1,795	1,676
Transition Markers	45	45	18,114	17,617

Table 4

Occurrences of Frame Markers and Transition Markers in NS and NNS Corpora ***

Corpus	Overall Tokens		Tokens at Sentence Initial Position	
	NS	NNS	NS	NNS
Frame M.	1,817.2 (709)	1,463.6 (549)	451.1 (176)	335.9 (126)
Transition M.	5,682.2 (2,217)	5,654.3 (2,121)	2,788.6 (1,088)	3,377.6 (1,267)

Note. Only markers that appeared at the sentence-initial position were included in this table. *Per million tokens are used as the unit. **Actual frequency is in parentheses.

Since interactive metadiscourse markers are characteristically used to signal moves and steps in the discourse, we further analyzed the interactive metadiscourse markers in more detail by running concordance software for each and every metadiscourse marker we identified. After that, we chose the markers that were used at the beginning of a sentence. As there were significantly more types of frame and transition markers and also as previous

research suggests the importance of these two types of metadiscourse, we focused on these two types. Table 4 summarizes the frequencies of frame markers and transition markers that appeared in both corpora. While NNS used more transition markers at the beginning of the sentence, NS constantly used more frame markers.

Table 5.

Comparison of Frame Markers between NS Corpus and NNS Corpus per Million Words

Frame M.	*NS*	*NNS*	*+/−*
Sequencing			
finally	174.3	98.6	−75.7
first	51.3	48	−3.3
firstly	2.6	2.7	+0.1
lastly	7.7	0	−7.7
second	33.3	29.3	−4.0
secondly	2.6	2.7	+0.1
then	10.3	42.7	+32.4
third	10.3	10.7	+0.4
Label Stages			
in conclusion	43.6	58.6	+15.0
in sum	2.6	0	−2.6
in summary	35.9	10.7	−25.2
overall	69.2	26.7	−42.5
so far	7.7	5.3	−2.4

Table 6

Usage of Transition Markers by NS and NNS per Million Words

Corpus	*NS*	*NNS*	*+/−*
Comparison/Contrast			
although	358.8	415.9	+57.1
in contrast	310.1	367.9	+57.8
while	153.8	53.3	−100.5
conversely	69.2	29.3	−39.9
by contrast	58.9	61.3	+2.4
whereas	41.0	32.0	−9.0
likewise	25.6	8.0	−17.6
yet	25.6	10.7	−14.9
nevertheless	23.1	10.7	−12.4
on the other hand	20.5	125.3	+104.8
Cause/Effect			
because	89.7	93.3	+3.6
since	76.9	189.3	+112.4
Conclusion			
thus	353.7	306.6	−47.1
therefore	89.7	277.2	+187.5
so	38.4	10.7	−27.7
consequently	33.3	26.7	−6.6
hence	28.2	21.3	−6.9
accordingly	20.5	8.0	−12.5
as a result	12.8	37.3	+24.5
Addition			
furthermore	297.3	546.5	+249.2
in addition	269.1	383.9	+114.8
moreover	243.5	247.9	+4.4
additionally	76.9	26.7	−50.2

Table 5 shows all the frame markers with their frequency information. Although the NS corpus contains more frame markers in general, NNS used some phrases more often than NS (e.g., *then* and *in conclusion*) and some phrases significantly less often (e.g., *finally*, *in summary*, and *overall*).

Table 6 describes frequently used transition markers based on the usage, and this table also shows some overused and underused phrases by NNS compared to the NS norm.

Discussion

We found various types of metadiscourse markers in our data. Among them, interactive metadiscourse markers are used frequently to signal moves and steps in scientific abstracts by both NS and NNS. Among the metadiscourse markers, NS used frame markers that signal sequences more specifically (i.e., the order of events), whereas NNS used more general expressions such as the use of *then* instead of using *finally*.

As for transition markers, based on their use in discourse, we grouped them into four categories: comparison/contrast, cause/effect, conclusion, and addition. A close look at each category reveals specific characteristics of abstracts written by NNS. First of all, NNS used more transition markers, and the distribution for each marker varies between NS and NNS. We surmised there are mainly two reasons.

First, there is a grammatical factor in the use of transition markers. For example, the occurrences of adverbs ending with a morpheme "-ly" used at the beginning of a sentence (e.g., *accordingly, additionally, consequently*, etc.) were much fewer in the NNS corpus than in the NS corpus. Instead, NNS used more grammatically explicit conjunctions (e.g., *therefore, furthermore, moreover, in addition*, etc.). Likewise, we found NNS did not use conjunctions that require more complex grammatical structures when grammatically simpler structures are available (e.g., *on the other hand* instead of *while*).

Second, familiarity of vocabulary is another factor. While NNS used more familiar items, which are most likely taught in junior or senior high schools in Japan (e.g., *then, therefore*, and *furthermore* instead of *hence and likewise*), the use of adverbial phrases by NS was more varied. NNS scientists seem to prefer frequently used linguistic patterns rather than avoiding overuse of certain expressions for rhetorical reasons.

Our data show that even though there are some preferences of use, the overall pattern of metadiscourse use is shared by both NNS and NS scientists, suggesting NNS scientists follow the genre and discipline-specific style of discourse. However, their usage of such markers may present some characteristics of L2 writing.

In order to comprehend scientific abstracts, it would be helpful for L2 readers to know the patterns of frame markers and transition markers identified in our corpora, and most of metadiscourse markers are familiar items even for high school students in Japan. However, some specific constructions such as the use of adverbs as connectors can be taught more explicitly in scientific English courses, as their uses seem to be less familiar to Japanese English learners. For the instruction of writing abstracts in life sciences, these metadiscourse resources should be emphasized more to raise the awareness of learners.

References

AMA Manual of Style. American Medical Association. (2016, December 1). Retrieved from http://www.amamanualofstyle.com/view/10.1093/jama/9780195176339.001.0001/med-9780195176339-div1-21

Bhatia, V. K. (1993). *Analysing genre: Language use in professional settings*. London: Longman.

Cao, F. & Hu, G. (2014). Interactive metadiscourse in research articles: A comparative study of paradigmatic and disciplinary influences. *Journal of Pragmatics 66*, 15-31.

Carrió-Pastor, M. (2013). A contrastive study of the variation of sentence connectors in academic English. *Journal of English for Academic Purposes 12,* 192-202.

Connor, U. & Mauranen, A. (1999). Linguistic analysis of grant proposals: European Union research grants. *English for Specific Purposes 18*(1), 47-62.

Cross, C. & Oppenheim, C. (2006). A genre analysis of scientific abstracts. *Journal of Documentation 62*(4), 428 – 446.

Dahl, T. (2004). Textual metadiscourse in research articles: a marker of national culture or of academic discipline? *Journal of Pragmatics. 36*. 1807-1825.

Hyland, K. & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics 25*(2). 156-177.

Hyland, K (2005). *Metadiscourse: Exploring interaction in writing*. London: Continuum.

Nature (2016, December 1). Nature guide to authors: First paragraphs for letters. Retrieved from www.nature.com/nature/authors/gta

Peacock, M. (2010). Linking adverbials in research articles across eight disciplines. *Iberica 20*. 9-34.

Samraj, B. (2005). An exploration of a genre set: Research article abstracts and introductions in two disciplines. *English for Specific Purposes 24*. 141-156.

Stoller, F.L. & Robinson, M.S. (2013). Chemistry journal articles: An interdisciplinary approach to move analysis with pedagogical aims. *English for Specific Purposes 32*. 45-57.